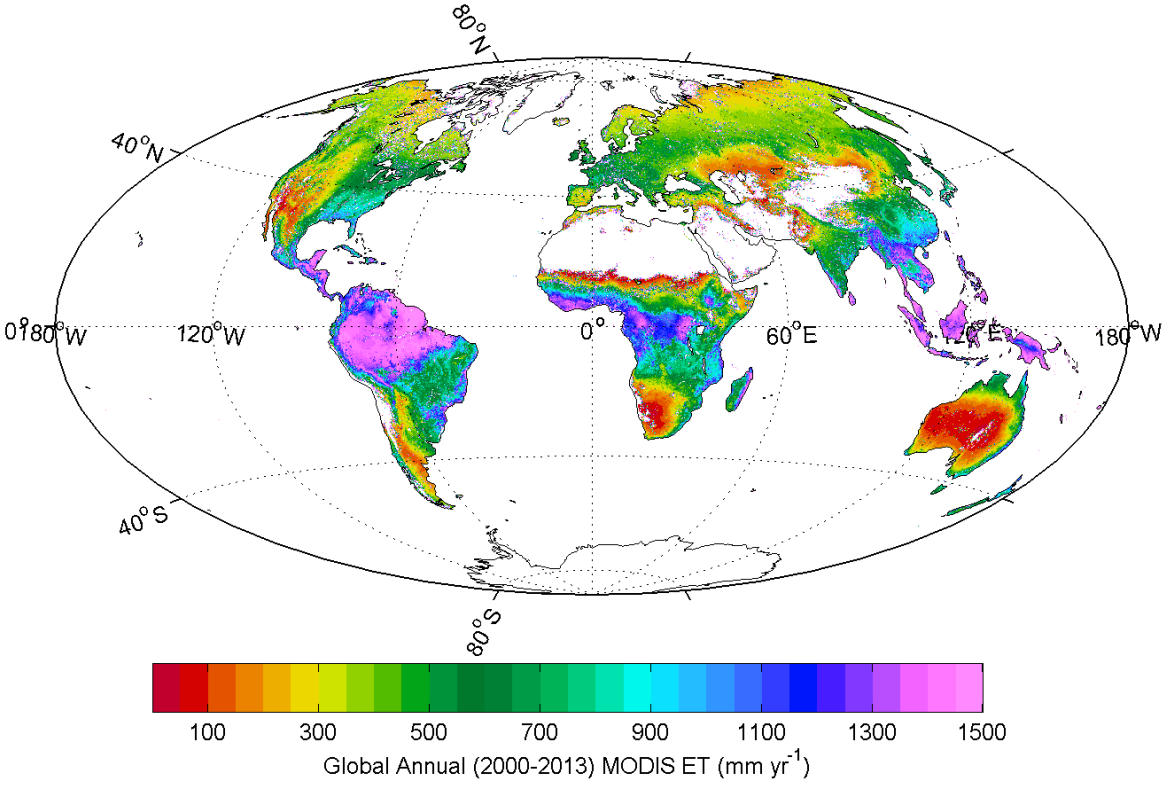
User's Guide
MODIS Global Terrestrial Evapotranspiration (ET) Product
(NASA MOD16A2/A3)
NASA Earth Observing System MODIS Land Algorithm

Steven W. Running
Qiaozhen Mu
Maosheng Zhao
Alvaro Moreno



Version 1.5
For Collection 6
May 5th, 2017

Table of Contents

Synopsis

This user's guide describes a level 4 MODIS land data product, MOD16, the global 8-day (MOD16A2) and annual (MOD16A3) terrestrial ecosystem Evapotranspiration (ET) dataset at 0.5 km spatial resolution over the 109.03 Million km2 global vegetated land areas designed for the MODIS sensor aboard the Aqua and Terra platforms. The MOD16 algorithm is based on the logic of the Penman-Monteith equation which uses daily meteorological reanalysis data and 8-day remotely sensed vegetation property dynamics from MODIS as inputs. This document is intended to provide both a broad overview and sufficient detail to allow for the successful use of the data in research and applications.

Please note the "MOD" prefix should be considered as referring to data sets derived from MODIS onboard either TERRA or Aqua satellite. That is, "MOD" in this document can also be treated as "MYD" derived from MODIS on Aqua.

1. The Algorithm, Background and Overview

Calculation of ET is typically based on the conservation of either energy or mass, or both. Computing ET is a combination of two complicated major issues: (1) estimating the stomatal conductance to derive transpiration from plant surfaces; and (2) estimating evaporation from the ground surface. The MOD16 ET algorithm runs at daily basis and temporally, daily ET is the sum of ET from daytime and night. Vertically, ET is the sum of water vapor fluxes from soil evaporation, wet canopy evaporation and plant transpiration at dry canopy surface. Remote sensing has long been recognized as the most feasible means to provide spatially distributed regional ET information on land surfaces. Remotely sensed data, especially those from polar-orbiting satellites, provide temporally and spatially continuous information over vegetated surfaces useful for regional measurement and monitoring of surface biophysical variables affecting ET, including albedo, biome type and leaf area index (LAI) (Los et al., 2000).

1.1 Energy Partitioning Logic

Energy partitioning at the surface of the earth is governed by the following three coupled equations:

$$H = \rho C_p \frac{T_s - T_a}{r_a} \quad (1)$$

$$\lambda E = \frac{\rho C_p (e_{sat} - e)}{\lambda (r_a + r_s)} \quad (2)$$

$$A' = R_{net} - \Delta S - G = H + \lambda E \quad (3)$$

where H, λE and A' are the fluxes of sensible heat, latent heat and available energy for H and λE; R_{net} is net radiation, G is soil heat flux; ΔS is the heat storage flux. λ is the latent heat of vaporization. ρ is air density, and C_P is the specific heat capacity of air; T_s, T_a are the aerodynamic surface and air temperatures; r_a is the aerodynamic resistance; e_{sat}, e are the water

vapour pressure at the evaporating surface and in the air; r_s is the surface resistance to evapotranspiration, which is an effective resistance to evaporation from land surface and transpiration from the plant canopy. The psychrometric constant γ is given by

$$\gamma = \frac{C_p P_a M_a}{\lambda M_w} \quad (4)$$

where M_a and M_w are the molecular masses of dry air and wet air respectively and P_a the atmospheric pressure.

1.2 Penman-Monteith Logic

Developing a robust algorithm to estimate global evapotranspiration is a significant challenge. Traditional energy balance models of ET require explicit characterization of numerous physical parameters, many of which are difficult to determine globally. For these models, thermal remote sensing data (e.g., land surface temperature, LST) are the most important inputs. However, using the 8-day composite MODIS LST (the average LST of all cloud-free data in the compositing window) (Wan et al., 2002) and daily meteorological data recorded at the flux tower, Cleugh et al. (2007) demonstrate that the results from thermal models are unreliable at two Australian sites (Virginia Park, a wet/dry tropical savanna located in northern Queensland and Tumbarumba, a cool temperate, broadleaved forest in south east New South Wales). Using a combination of remote sensing and global meteorological data, we have adapted the Cleugh et al. (2007) algorithm, which is based on the Penman–Monteith method and calculates both canopy conductance and ET. Monteith (1965) eliminated surface temperature from Equations (1) – (3) to give:

$$\lambda E = \frac{s A' + \rho C_p \dfrac{(e_{sat} - e)}{r_a}}{s + \gamma \left(1 + \dfrac{r_s}{r_a}\right)} = \frac{s A' + \rho C_p \dfrac{VPD}{r_a}}{s + \gamma \left(1 + \dfrac{r_s}{r_a}\right)} \quad (5)$$

where $s = d(e_{sat})/dT$, the slope of the curve relating saturated water vapor pressure (e_{sat}) to temperature; A' is available energy partitioned between sensible heat and latent heat fluxes on land surface. $VPD = e_{sat} - e$ is the air vapor pressure deficit. All inputs have been previously defined except for surface resistance r_s, which is an effective resistance accounting for evaporation from the soil surface and transpiration from the plant canopy.

Despite its theoretical appeal, the routine implementation of the P-M equation is often hindered by requiring meteorological forcing data (A', T_a and VPD) and the aerodynamic and surface resistances (r_a and r_s). Radiation and soil heat flux measurements are needed to determine A'; air temperature and humidity to calculate VPD; and wind speed and surface roughness parameters to determine r_a. Multi-temporal implementation of the P-M model at regional scales requires routine surface meteorological observations of air temperature, humidity, solar radiation and wind speed. Models for estimating maximum stomatal conductance including the effect of limited soil water availability and stomatal physiology requires either a fully coupled

biophysical model such as that by Tuzet et al. (2003) or resorting to the empirical discount functions of Jarvis (1976), which must be calibrated. Determining a surface resistance for partial canopy cover is even more challenging with various dual source models proposed (e.g. Shuttleworth and Wallace, 1985) to account for the presence of plants and soil.

2. The MOD16A2/MOD16A3 algorithm logic

MOD16 ET algorithm is based on the Penman-Monteith equation (Monteith, 1965) as in equation 5. Figure 2 shows the logic behind the improved MOD16 ET Algorithm for calculating daily MOD16 ET algorithm.

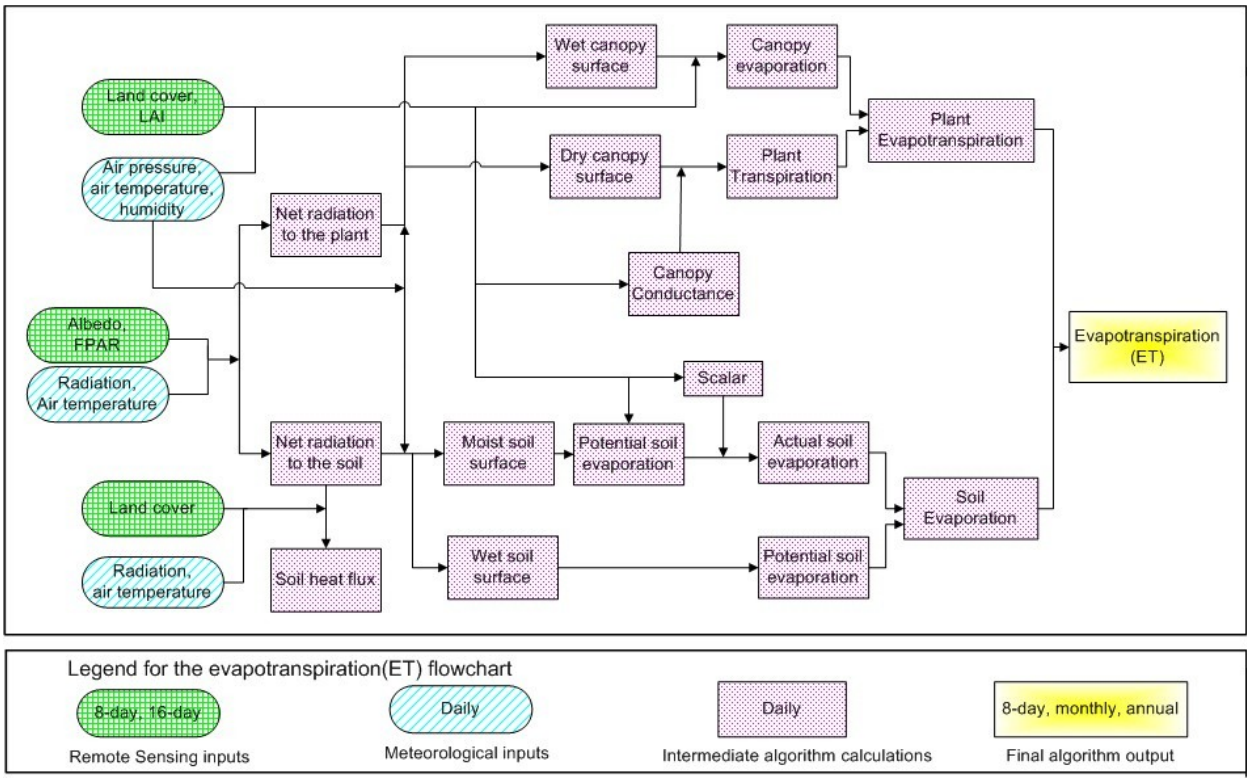


Figure 2. Flowchart of the improved MOD16 ET algorithm. LAI: leaf area index; FPAR: Fraction of Photosynthetically Active Radiation.

2.1. Daytime and Nighttime ET

Daily ET should be the sum of daytime and nighttime ET. To get nighttime average air temperature (T_{night}), we assume that daily average air temperature (T_{avg}) is the average of daytime air temperature (T_{day}) and T_{night}.

$$T_{night} = 2\,T_{day} - T_{night} \qquad (6)$$

The net incoming solar radiation at night is assumed to be zero. Based on the optimization theory, stomata will close at night to prevent water loss when there is no opportunity for carbon gain (Dawson et al., 2007). In the improved ET algorithm, at night, the stomata are

assumed to close completely and the plant transpiration through stomata is zero, except for the transpiration through leaf boundary-layer and leaf cuticles (more details in section 2.5). Both nighttime and daytime use the same ET algorithm except that different values at daytime and nighttime are used for the same variable.

2.2. Soil Heat Flux

In MOD16 ET algorithm, the net incoming radiation to the land surface (R_{net}) is calculated as the equations 7 and 8 (Cleugh et al., 2007).

$$R_{net} = (1-a)R_{s\downarrow} + \sigma(\varepsilon_a - \varepsilon_s)\left(273.15+T\right)^4 \quad (7)$$

$$\varepsilon_a = 1 - 0.26\, e^{(7.77 \cdot 10^{-4} T^2)}$$

$$\varepsilon_s = 0.97$$

where α is MODIS albedo, $R_{s\downarrow}$ is the downward shortwave radiation, ε_s is surface emissivity, ε_a is atmospheric emissivity, and T is air temperature in °C. At daytime, if R_{net} is less than zero, R_{net} is set to be zero; at nighttime, if R_{net} is less than -0.5 times of daytime R_{net}, nighttime R_{net} is set as -0.5 multiplying daytime R_{net}. There is no soil heat flux (G) interaction between the soil and atmosphere if the ground is 100% covered with vegetation. Energy received by soil is the difference between the radiation partitioned on the soil surface and soil heat flux (G).

$$A = R_{net} \quad (8)$$

$$A_c = F_c A$$

$$A_{soil} = (1-F_c)A - G$$

where A is available energy partitioned between sensible heat, latent heat and soil heat fluxes on land surface; R_{net} is the net incoming radiation received by land surface; A_c is the part of A allocated to the canopy and A_{SOIL} is the part of A partitioned on the soil surface. Net radiation is partitioned between the canopy and soil surface based on vegetation cover fraction (FC), in order to reduce numbers of inputs from MODIS datasets and to simplify the algorithm, we use 8-day 0.5 km2 MOD15A2H FPAR (the Fraction of Absorbed Photosynthetically Active Radiation) as a surrogate of vegetation cover fraction (Los et al., 2000), $FC = FPAR$.

At the extremely hot or cold places or when the difference between daytime and nighttime temperature is low (<5°C), there is no soil heat flux. The soil heat flux is estimated as:

$$G_{soil}=\begin{cases} 4.73\,T_i-20.87\;T_{minclose}\le Tann_{avg}<25℃\,,\,T_{day}-T_{night}\ge 5℃ \\ 0\;Tann_{avg}\ge 25℃\vee Tann_{avg}<T_{minclose}\vee T_{day}-T_{night}\ge 5℃ \\ 0.39\;A_i\left|\left(G_{soil}\right)\right|>0.39\left|\left(A_i\right)\right| \end{cases}$$

$$G=G_{soil}\left(1-F_c\right) \quad (9)$$

where G_{soil} stands for the soil heat flux when $F_C = 0$; $Tann_{avg}$ is annual average daily temperature, and $T_{min\ close}$ is the threshold value below which the stomata will close completely and halt plant transpiration (Table 3.2; Running et al., 2004; Mu et al., 2007; Mu et al., 2011). At daytime, $G_{soil\ day} = 0.0$ if $A_{day} - G_{soil\ day} < 0.0$; at nighttime, $G_{soil\ night} = A_{night} + 0.5A_{day}$ if $A_{day} > 0.0$ and $A_{night} - G_{night} < -0.5\ A_{day}$.

2.3. Wet Surface Fraction

In the MOD16 algorithm, ET is the sum of water lost to the atmosphere from the soil surface through evaporation, canopy evaporation from the water intercepted by the canopy, and transpiration from plant tissues (Fig. 2). The land surface is covered by the plant and the bare soil surface, and percentage of the two components is determined by F_C. Both the canopy and the soil surface are partly covered by water under certain conditions. The water cover fraction (F_{wet}) is taken from the Fisher et al. (2008) ET model, modified to be constrained to zero when relative humidity (RH) is less than 70%:

$$F_{wet}=\begin{cases} 0\;RH\le 100 \\ RH^4\;70\le RH\le 100 \end{cases} \quad (10)$$

where RH is relative humidity (Fisher et al, 2008). When RH is less than 70%, 0% of the surface is covered by water. For the wet canopy and wet soil surface, the water evaporation is calculated as the potential evaporation as described in the next sections (2.4 and 2.6).

2.4. Evaporation from Wet Canopy Surface

Evaporation of precipitation intercepted by the canopy accounts for a substantial amount of upward water flux in ecosystems with high LAI. When the vegetation is covered with water (i.e., F_{wet} is not zero), water evaporation from the wet canopy surface will occur. ET from the vegetation consists of the evaporation from the wet canopy surface and transpiration from plant tissue, whose rates are regulated by aerodynamics resistance and surface resistance.

The aerodynamic resistance ($rhrc$, s m^{-1}) and wet canopy resistance (rvc, s m^{-1}) to evaporated water on the wet canopy surface are calculated as

$$rhc = \frac{1}{g\, l_{sh}\, LAI\, F_{wet}} \quad (11)$$

$$rrc = \frac{\rho\, C_p}{4\sigma (T_i + 273.15)^4}$$

$$rhrc = \frac{rhc\, rrh}{rhc + rrh}$$

$$rvc = \frac{1}{g\, l_{ewv}\, LAI\, F_{wet}}$$

where *rhc* (s m^{-1}) is the wet canopy resistance to sensible heat, *rrc* (s m^{-1}) is the resistance to radiative heat transfer through air; gl_{sh} (s m^{-1}) is leaf conductance to sensible heat per unit LAI, $gl_{e\ wv}$ (m s^{-1}) is leaf conductance to evaporated water vapor per unit LAI, σ (W m^{-2} K^{-4}) is Stefan-Boltzmann constant.

Following Biome-BGC model (Thornton, 1998) with revision to account for wet canopy, the evaporation on wet canopy surface is calculated as

$$\lambda E_{wet\,c} = \frac{\left(s\, A_c + \rho\, C_p\, F_c \frac{(e_{sat} - e)}{rhrc} \right) F_{wet}}{s + \frac{P_a C_p\, rvc}{\lambda\, \varepsilon\, rhrc}} \quad (12)$$

where the resistance to latent heat transfer (*rvc*) is the sum of aerodynamic resistance (*rhrc*) and surface resistance (r_s) in equation 5.

2.5. Plant Transpiration

2.5.1. Surface Conductance to Transpiration

Plant transpiration occurs not only during daytime but also at nighttime. For many plant species, stomatal conductance (G) decreases as vapor pressure deficit (VPD) increases, and stomatal conductance is further limited by both low and high temperatures (Jarvis, 1976; Sandford et al., 1986; Kawamitsu et al., 1993; Schulze et al., 1994; Leuning, 1995; Marsden et al., 1996; Dang et al., 1997; Oren et al., 1999; Xu et al., 2003; Misson et al., 2004). Because high temperatures are often accompanied by high VPDs, we have only added constraints on stomatal conductance for VPD and minimum air temperature ignoring constraints resulting from

high temperature. Based on the optimization theory, stomata will close at night to prevent water loss when there is no opportunity for carbon gain (Dawson et al., 2007).

$$G_{si}^{1}=\begin{cases} C_L m(Tmin) m(VPD) r_{corr} & i=daytime \\ 0 & i=nighttime \end{cases}$$

$$r_{corr}=\frac{1}{\frac{101300}{P_a}\left(\frac{T_i+273.15}{293.15}\right)^{1.75}} \quad (13)$$

$$m(T_{min})=\begin{cases} 1 & T_{min}\geq T_{min\,open} \\ \frac{T_{min}-T_{min\,close}}{T_{min\,open}-T_{min\,close}} & T_{min\,close}<T_{min}<T_{min\,open} \\ 0 & T_{min}\leq T_{min\,close} \end{cases}$$

$$m(VPD)=\begin{cases} 1 & VPD\leq VPD_{open} \\ \frac{VPD_{close}-VPD}{VPD_{close}-VPD_{open}} & VPD_{open}<VPD<VPD_{close} \\ 0 & VPD\geq VPD_{close} \end{cases}$$

where C_L is the mean potential stomatal conductance per unit leaf area, C_L is set differently for different biomes as shown in Table 3.2 (Kelliher et al., 1995; Schulze et al., 1994; White et al., 2000), $m(T_{min})$ is a multiplier that limits potential stomatal conductance by minimum air temperatures (T_{min}), and $m(VPD)$ is a multiplier used to reduce the potential stomatal conductance when VPD (difference between e_{sat} and e) is high enough to reduce canopy conductance. Sub index *close* indicates nearly complete inhibition (full stomatal closure) due to low T_{min} and high VPD, and *open* indicates no inhibition to transpiration (Table 3.2). When T_{min} is lower than the threshold value $T_{min\ close}$, or VPD is higher than the threshold VPD_{close}, the strong stresses from temperature or water availability will cause stomata to close completely, halting plant transpiration. On the other hand, when T_{min} is higher than $T_{min\ open}$, and VPD is lower than VPD_{open}, there will be no temperature or water stress on transpiration. For T_{min} and VPD falling into the range of the upper and low limits, the corresponding multiplier will be within 0.0 to 1.0, implying a partial stomatal closure. The multipliers range linearly from 0 (total inhibition, limiting r_S) to 1 (no inhibition) for the range of biomes are listed in a Biome Properties Look-Up Table (BPLUT) (Table 3.2).

The reason to use the correction function r_{corr} is that, the conductance through air varies with the air temperature and pressure. The prescribed values are assumed to be given for standard conditions of 20˚C and 101300 Pa. Based on the prescribed daily air temperature (converted to Kelvins) and an air pressure estimated from a prescribed elevation, the prescribed standard conductances are converted to actual conductances for the day according to Jones (1992) and Thornton (1998). P_a is calculated as a function of the elevation (Thornton, 1998).

$$t_1 = 1 - \frac{LR_{STD} Elev}{T_{STD}}$$

$$t_2 = 1 - \frac{G_{STD}}{LR_{STD} \frac{RR}{MA}} \quad (14)$$

$$P_a = P_{STD} t_1^{t_2}$$

where LR_{STD}, T_{STD}, G_{STD}, RR, MA and P_{STD} are constant values as listed in Table 1.1 LR_{STD} (K m^{-1}) is standard temperature lapse rate; $TSTD$ (K) is standard temperature at 0.0 m elevation; G_{STD} (m s^{-2}) is standard gravitational acceleration; RR (m^3 P_a mol^{-1} K^{-1}) is gas law constant; MA (kg mol^{-1}) is molecular weight of air and $PSTD$ (Pa) is standard pressure at 0 m elevation.

Table 1.1 Other parameter values as used in the improved ET algorithm

LR_{STD} ($K\ m^{-1}$)	T_{STD} (K)	G_{STD} ($m\ s^{-2}$)	RR ($m^3\ Pa\ mol^{-1}\ K^{-1}$)	MA ($kg\ mol^{-1}$)	P_{STD} (Pa)
0.0065	288.15	9.80665	8.3143	28.9644e-3	101325.0

Canopy conductance (C_c) to transpired water vapor per unit LAI is derived from stomatal and cuticular conductances in parallel with each other, and both in series with leaf boundary layer conductance (Thornton, 1998; Running & Kimball, 2005). In the case of plant transpiration, surface conductance is equivalent to the canopy conductance (C_C), and hence surface resistance (r_s) is the inverse of canopy conductance (C_C).

$$C_{ci} = \begin{cases} \frac{G_s^2 (G_{si}^1 + G_{CU})}{G_{si}^1 + G_s^2 + G_{CU}} (LAI > 0, (1 - F_{wet}) > 0) \\ 0 (LAI = 0, (1 - F_{wet}) = 0) \end{cases} \quad (15)$$

$$G_{cu} = g_{cu} r_{corr}$$

$$G_s^2 = gl_{sh}$$

$$r_{si} = \frac{1}{C_{ci}}$$

where the subscript i means the variable value at daytime and nighttime; G_{CU} is leaf cuticular conductance; G_s^2 is leaf boundary-layer conductance; g_{cu} is cuticular conductance per unit LAI, set as a constant value of 0.00001 (m s^{-1}) for all biomes; gl_{sh} is leaf conductance to sensible heat per unit LAI, which is a constant value for each given biome (Table 3.2).

2.5.2. Aerodynamic Resistance

The transfer of heat and water vapor from the dry canopy surface into the air above the canopy is determined by the aerodynamic resistance (r_a). r_a is calculated as a parallel resistance to convective (rh) and radiative (rr) heat transfer following Biome-BGC model (Thornton, 1998).

$$r_a = \frac{rh\,rr}{rh+rr}$$

$$rh = \frac{1}{gl_{bl}} \quad (16)$$

$$rr = \frac{\rho C_p}{4\sigma (T_i + 273.15)^3}$$

where gl_{bl} (m s^{-1}) is leaf-scale boundary layer conductance, whose value is equal to leaf conductance to sensible heat per unit LAI (gl_{sh} (m s^{-1}) as in section 2.4), and σ (W m^{-2} K^{-4}) is Stefan-Boltzmann constant.

2.5.3. Plant Transpiration

Finally, the plant transpiration (AE_{trans}) is calculated as

$$\left(s A_c + \rho C_p F_c \frac{(e_{sat} - e)}{r_a} \right) \frac{(¿¿wet)}{(1 \quad r_s)} \quad (17)$$

where r_a is the aerodynamic resistance calculated from equation 5.

$$\alpha s A_c \frac{(¿¿wet)}{} \quad (18)$$

with α = 1.26.

2.6. Evaporation from Soil Surface

The soil surface is divided into the saturated surface covered with water and the moist surface by F_{wet}. The soil evaporation includes the potential evaporation from the saturated soil surface and evaporation from the moist soil surface. The total aerodynamic resistance to vapor transport (r_{tot}) is the sum of surface resistance (r_s) and the aerodynamic resistance for vapor transport (r_v) such that $r_{tot} = r_v + r_s$ (van de Griend, 1994; Mu et al., 2007). We assume that r_v (*s m-1*) is equal to the aerodynamic resistance (r_a: *s m-1*) in Equation 5 since the values of r_v and *ra* are usually very close (van de Griend, 1994). In the MOD16 ET algorithm, the *rhs* is assumed to be equal to boundary layer resistance, which is calculated in the same way as total aerodynamic resistance (*rtot*) (Thornton, 1998) only that, *rtotc* is not a constant. For a given biome type, there is a maximum (*rblmax*) and a minimum value (*rblmin*) for *rtotc*, and *rtotc* is a function of VPD.

$$r_{tot} = r_{totc} r_{corr} \quad (19)$$

$$r_{totc} = \left\{ rbl_{max} - \frac{(rbl_{max} - rbl_{min})(VPD_{close} - VPD)}{VPD_{close} - VPD_{open}} \right. \quad VPD_{open} < VPD < VPD_{close}$$

where *rcorr* is the correction for atmospheric temperature and pressure (equation 13) above mentioned. The values of *rblmax* and *rblmin*, *VPDopen* (when there is no water stress on transpiration) and *VPDclose* (when water stress causes stomata to close almost completely, halting plant transpiration) are parameterized differently for different biomes and are listed in Table 3.2.

The aerodynamic resistance at the soil surface (r_{as}) is parallel to both the resistance to convective heat transfer (*rhs*: s m-1) and the resistance to radiative heat transfer (r_{rs}: s m-1) (Choudhury and DiGirolamo, 1998), such that

$$r_{as} = \frac{r_{hs} r_{rs}}{r_{hs} + r_{rs}}$$

$$r_{hs} = r_{tot}(20)$$

$$r_{rs} = \frac{\rho C_p}{4 \sigma (T_a + 273.15)^3}$$

The actual soil evaporation (λE_{SOIL}) is calculated in equation 34 using potential soil evaporation ($\lambda E_{SOILPOT}$) and soil moisture constraint function in the Fisher et al. (2008) ET model. This function is based on the complementary hypothesis (Bouchet, 1963), which defines land-atmosphere interactions from air *VPD* and relative humidity (*RH*, %).

$$\lambda E_{wet\,soil} = \frac{\left(s A_{soil} + \rho C_p (1 - F_c) \frac{VPD}{r_{as}} \right) F_{wet}}{s + \gamma \left(1 + \frac{r_{ss}}{r_{as}} \right)}$$

$$\lambda E_{soil} = \frac{\left(s A_{soil} + \rho C_p (1 - F_c) \frac{VPD}{r_{as}} \right) (1 - F_{wet})}{s + \gamma \left(1 + \frac{r_{ss}}{r_{as}} \right)} \quad (21)$$

$$\lambda E_{soil} = \lambda E_{wet\,soil} + \lambda E_{soil} \left(\frac{RH}{100} \right)^{\frac{VPD}{\beta}}$$

with β = 200.

2.7. Total Daily Evapotranspiration

The total daily ET is the sum of evaporation from the wet canopy surface, the transpiration from the dry canopy surface and the evaporation from the soil surface. The total daily ET and potential ET*(λE_{POT})* are calculated as in equation 22.

$$\lambda E = \lambda E_{wet\,c} + \lambda E_{trans} + \lambda E_{soil} \quad (22)$$

$$\lambda E_{pot} = \lambda E_{wet\,c} + \lambda E_{pot\,trans} + \lambda E_{wet\,soil} + \lambda E_{soil\,pot}$$

Combination of ET with the potential ET can determine environmental water stress and detect the intensity of drought.

2.8. Updates after Publication of RSE Paper by Mu et al. (2011)

The MOD16 products are generated based on the MOD16 algorithm in Mu et al.'s 2011 RSE paper. Since the publication, Dr. Mu have updated the product to fix some issues and these updates have been implemented in the operational code.

1. There were negative ET and even PET values for some 8-day and monthly data in the. In the code for the previous product, we allowed the net incoming daytime radiation to be negative. Only MERRA daytime downward solar radiation, daytime actual vapor pressure, daytime temperature, daily average and minimum temperatures are used as meteorological inputs. The outgoing and incoming longwave radiation is calculated as in Mu et al.'s 2011 RSE paper. In the updated product, the nighttime actual vapor pressure, nighttime temperature, outgoing and incoming longwave radiation are from MERRA directly. When the daytime net incoming radiation is negative, 0 will be used instead.

In the MOD16 ET algorithm, when we calculated the soil heat flux, we didn't constrain the soil heat flux. The net radiation to the bare ground is difference between the fraction of the net incoming radiation reaching to the ground surface (Rground) and the soil heat flux (G). In certain cases, G is higher than Rground, resulting in negative soil evaporation. In the updated product, We set limit to G. If ((daytime net incoming radiation) - (daytime G)) < 0, then we set (daytime G)=(daytime net incoming radiation). If ((nighttime net incoming radiation) - (nighttime G)) < -0.5*(daytime net incoming radiation), then we set (nighttime G)=(nighttime net incoming radiation) + 0.5*(daytime net incoming radiation).

2. Filled value for vegetated pixels when there are no valid MODIS surface albedo values throughout the year due to high frequency of cloudiness in the improved MOD16. For example, this occurs over rainforests of west Africa due to severe and constant cloudiness. In the update version, we specify an albedo value of 0.4 for the pixels, a typical value for nearby rainforests with valid albedo values.

3. Operational Details of MOD16 and Primary Uncertainties in the MOD16 Logic

A discusses some of ables, and their infl

3

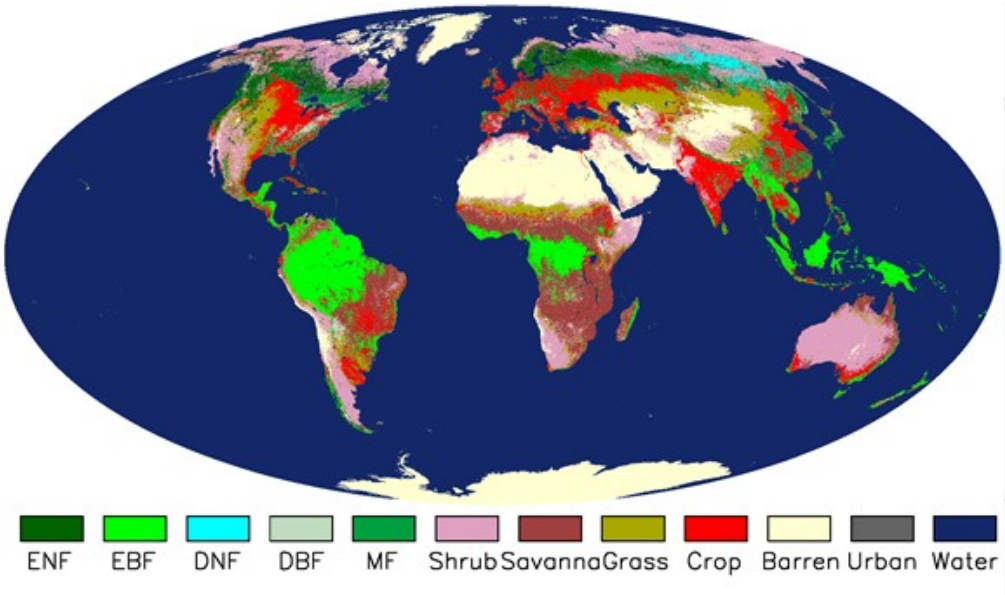


Figure 3.1. *University of Maryland land cover classification system defined land covers from MOD12Q1 (land cover type 2 in MOD12Q1 dataset). Evergreen Needleleaf Forest (ENF), Evergreen Broadleaf Forest (EBF), Deciduous Needleleaf Forest (DNF), Deciduous Broadleaf Forest (DBF), Mixed forests (MF), Closed Shrublands (CShrub), Open Shrublands (OShrub), Woody Savannas (WSavanna), Savannas (Savanna), grassland (Grass), and Croplands (Crop). Note that in this figure, we combined CShrub and OShrub into Shrub, and WSavanna and Savannas into Savanna.*

One of the first MODIS products used in the MOD16 algorithm is the Land Cover Product, MOD12Q1. The importance of this product cannot be overstated as the MOD16 algorithm relies heavily on land cover type through use of the BPLUT. While the primary product created by MOD12 is a 17-class IGBP (International Geosphere-Biosphere Programme) landcover classification map (Running et al. 1994, Belward et al. 1999, Friedl et al. 2010), the MOD16 algorithm employs Boston University's UMD classification scheme (Table 3.1). More details on these and other schemes and their quality control considerations can be found at: https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mcd12q1

3.1.1. The BPLUT and constant biome properties

Arguably, the most significant assumption made in the MOD16 logic is that biome-specific physiological parameters do not vary with space or time. These parameters are outlined in the BPLUT (Table 3.2) within the MOD16 algorithm. The BPLUT constitutes the physiological framework for controlling simulated carbon sequestration. These biome-specific properties are not differentiated for different expressions of a given biome, nor are they varied at any time during the year. In other words, a semi-desert grassland in Mongolia is treated the same as a tallgrass prairie in the Midwestern United States. Likewise, a sparsely vegetated boreal evergreen needleleaf forest in Canada is functionally equivalent to its coastal temperate evergreen needleleaf forest counterpart.

Table 3.1*. The land cover types used in the MOD16 Algorithm.*

UMD Land Cover Types	
Class Value	**Class Description**
0	Water
1	Evergreen Needleleaf Forest
2	Evergreen Broadleaf Forest
3	Deciduous Needleleaf Forest
4	Deciduous Broadleaf Forest
5	Mixed Forest
6	Closed Shrubland
7	Open Shrubland
8	Woody Savanna
9	Savanna
10	Grassland
12	Cropland
13	Urban or Built-Up
16	Barren or Sparsely Vegetated
254	Unclassified
255	Missing Data

Table 3.2. *Biome-Property-Look-Up-Table (BPLUT) for MODIS ET algorithm with NCEP-DOE reanalysis II and the Collection 6 FPAR/LAI/Albedo as inputs. The full names for the University of Maryland land cover classification system (UMD_VEG_LC) in MOD12Q1 dataset (fieldname: Land_Cover_Type_2) are, Evergreen Needleleaf Forest (ENF), Evergreen Broadleaf Forest (EBF), Deciduous Needleleaf Forest (DNF), Deciduous Broadleaf Forest (DBF), Mixed forests (MF), Closed Shrublands (CShrub), Open Shrublands (OShrub), Woody Savannas (WSavanna), Savannas (Savanna), Grassland (Grass), and Croplands (Crop). Note this table has been updated since publication of RSE paper by Mu et al. (2011), and please refer to section 2.8 for details.*

UMD_VEG_LC	**ENF**	**EBF**	**DNF**	**DBF**	**MF**	**CShrub**	**OShrub**	**WSavanna**	**Savanna**	**Grass**	**Crop**
$Tmin_{close}$ (C)	-8.00	-8.00	-8.00	-6.00	-7.00	-8.00	-8.00	-8.00	-8.00	-8.00	-8.00
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	11.39	11.39	12.02	12.02
VPD_{close} (Pa)	3000.0	4000.0	3500.0	290[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
gl_{sh} (m s^{-1})	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.04	0.04	0.02	0.02
$gl_{e\,wv}$ (m s^{-1})	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.04	0.04	0.02	0.02
g_{cu} (m s^{-1})	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
C_L (m s^{-1})	0.00240	0.00240	0.00240	0.00240	0.00240	0.00550	0.00550	0.00550	0.00550	0.00550	0.00550
rbl_{min} (s m^{-1})	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0
rbl_{max} (s m^{-1})	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0

3.2 Leaf area index, fraction of absorbed photosynthetically active radiation and albedo

The FPAR/LAI product is an 8-day composite product. The MOD15 compositing algorithm uses a simple selection rule whereby the maximum FPAR (across the eight days) is chosen for the inclusion as the output pixel. The same day chosen to represent the FPAR measure also contributes the current pixel's LAI value. This means that although ET is calculated daily, the MOD16 algorithm necessarily assumes that leaf area and FPAR do not vary during a given 8-day period. Compositing of LAI and FPAR is required to provide an accurate depiction of global leaf area dynamics with consideration of spectral cloud contamination, particularly in the tropics.

The MCD43A2/A3 albedo products are 8-day composite products too. Both Terra and Aqua data are used in the generation of this product, providing the highest probability for quality input data and designating it as an "MCD," meaning "Combined," product. Version-6 MODIS/Terra+Aqua BRDF/Albedo products are Validated Stage 1, indicating that accuracy has been estimated using a small number of independent measurements obtained from selected locations and time periods and ground-truth/field program efforts. Although there may be later improved versions, these data are ready for use in scientific publications.

3.2.1. Additional Cloud/Aerosol Screening

The 8-day MOD15A2H and MCD43A3 are still contaminated by clouds and/or aerosols in certain regions and times of year. As a result, in regions with higher frequencies of cloud cover, such as tropical rain forests, values of FPAR and LAI will be greatly reduced and the albedo signal dramatically increased. To distinguish between good quality and contaminated data, MOD15A2H and MCD43A3 contain Quality Control (QC) fields, enabling users to determine which pixels are suitable for further analysis. The use of contaminated FPAR, LAI or albedo inputs will produce incorrect ET estimates. Thus for near real-time NASA's operational MOD16 dataset, users should be cautious when use the data and contaminated MOD16 data sets should be excluded from analysis. In this case, QC data field should be used to distinguish uncontaminated or unreliable MOD16 pixels for the same data period. Only for the reprocessed and improved MOD16, users may ignore QC because the contaminated MODIS inputs to MOD16 have been temporally filled, resulting in an improved MOD16. The following describes the method we employed for the improved MOD16.

➶ *The following additional cloud/aerosol screening is* ***only applied to the improved MOD16*** *for the end of the year when the previous full yearly MOD15A2H/MCD43A3 are available (also see section 6.3). The screening processes for the improved MOD16* ***cannot be applied to NASA's near real-time operational MOD16*** *and users should exclude contaminated data sets based on QC data field. Users are advised to check PI's lab web page or USGS LP DAAC for information on the difference in the versions between the improved and operational MOD16.*

To improve the MOD16 data product, we solved this problem related to MOD15A2H/MCD43A3 inputs by removing poor quality FPAR/LAI/Albedo data based on the QC label for every pixel. If, for example, any LAI/FPAR pixel did not meet the quality screening criteria, its value is determined through linear interpolation between the previous period's value

and that of the next period to pass the screening process. Contaminated MOD15A2H/MCD43A3 was improved as the result of the filling process. However, there are some unusual 8-day periods with lower FPAR and LAI but good QC labels. In spite of this, we have to depend on QC label because this is the only source of quality control. Improved MOD15A2H/MCD43A3 leads to improvements of MOD16.

3.3. GMAO daily meteorological data

The MOD16 algorithm computes ET at a daily time step. This is made possible by the daily meteorological data, including average and minimum air temperature, incident PAR and specific humidity, provided by NASA's Global Modeling and Assimilation Office (GMAO or MERRA GMAO), a branch of NASA (Schubert et al. 1993). These data, produced every six hours, are derived using a global circulation model (GCM), which incorporates both ground and satellite-based observations. These data are distributed at a resolution of 0.5° x 0.6° (MERRA GMAO) or 1.00° x 1.25° (*note that resolution may become finer with updates of GMAO system at NASA*) in contrast to the 0.5 km gridded MOD16 outputs. It is assumed that the coarse resolution meteorological data provide an accurate depiction of ground conditions and are homogeneous within the spatial extent of each cell.

One major problem is the inconsistency in spatial resolution between half-degree GMAO/NASA meteorological data and 0.5 km MODIS pixel. We solved the problem by spatially smoothing meteorological data to 0.5 km MODIS pixel level. For the problem arising from coarse spatial resolution daily GMAO data, we use spatial interpolation to enhance meteorological inputs. The four GMAO cells nearest to a given 0.5 km MODIS pixel are used in the interpolation algorithm. There are two reasons for choosing four GMAO cells per 0.5 km MODIS pixel: (1) this will not slow down the computational efficiency of creating MOD16, which is a global product, and (2) it is more reasonable to assume no elevation variation within four GMAO cells than more GMAO cells.

Although there are many formulae for non-linear spatial interpolation, for simplicity, we use a cosine function because the output value can be constrained between 0 and 1. This function still could not effectively boundary lines in a MOD16 image, and thus we utilized a modified cosine function of the form:

$$D_i = \cos^4((\pi/2)(d_i/d_{\max})) \qquad i = 1,2,3,4 \tag{23}$$

where, D_i is the non-linear distance between the 0.5 km MODIS pixel and any one of four surrounding GMAO cells; d_i is the great-circle distance between the 0.5 km pixel and the same GMAO cell; and d_{max} is the great-circle distance between the two farthest GMAO cells of the four being used. This ensures that $D_i = 1$ when $d_i = 0$, and $D_i = 0$ when $d_i = d_{max}$.

Based on the non-linear distance (D_i), the weighted value W_i can be expressed as

$$W_i = D_i / \sum_{i=1}^{4} D_i \tag{24}$$

,

and therefore, for a given pixel, the corresponding smoothed value V (i.e., interpolated Tmin, Tavg, VPD, SWrad) is

$$V = \sum_{i=1}^{4} (W_i V_i) \quad (25)$$

Theoretically, this GMAO spatial interpolation can improve the accuracy of meteorological data for each 0.5 km pixel because it is unrealistic for meteorological data to abruptly change from one side of GMAO boundary to the other. To explore the above question, we use observed daily weather data from World Meteorological Organization (WMO) daily surface observation network (>5000 stations) to compare changes in Root Mean Squared Error (RMSE) and Correlation (COR) between the original and enhanced DAO data. As a result of the smoothing process, on average, RMSE is reduced and COR increased for 72.9% and 84% of the WMO stations, respectively, when comparing original and enhanced DAO data to WMO observations for 2001 and 2002. Clearly, the nonlinear spatial interpolation significantly improves GMAO inputs for most stations, although for a few stations, interpolated GMAO accuracy may be reduced due to the inaccuracy of GMAO in these regions. (Zhao et al. 2005, 2006)

4. Validation of MOD16

To validate the MOD16 algorithm, we used the observed latent heat flux for 46 field-based eddy covariance flux towers, global 232 watersheds, as well as global results over the past 11 years (2000 to 2010). We cut out the input MODIS data for the 3 x 3 1-km^2 pixels surrounding each tower. We drove the MOD16 ET algorithm with both tower observed meteorological data and global GMAO meteorological data. We got the average ET estimates over those of the 3 x 3 1-km^2 pixels where the tower actual vegetation type is the same as MOD12 land cover type 2. Then we compared the ET estimates with the tower ET observations. For each of the seven biome types among the 46 flux towers except for CSH and OSH since there is only one tower with fewer than 365 measurements for each of them, we chose one tower to show the performance of MOD16 ET algorithm (Fig. 4.1).

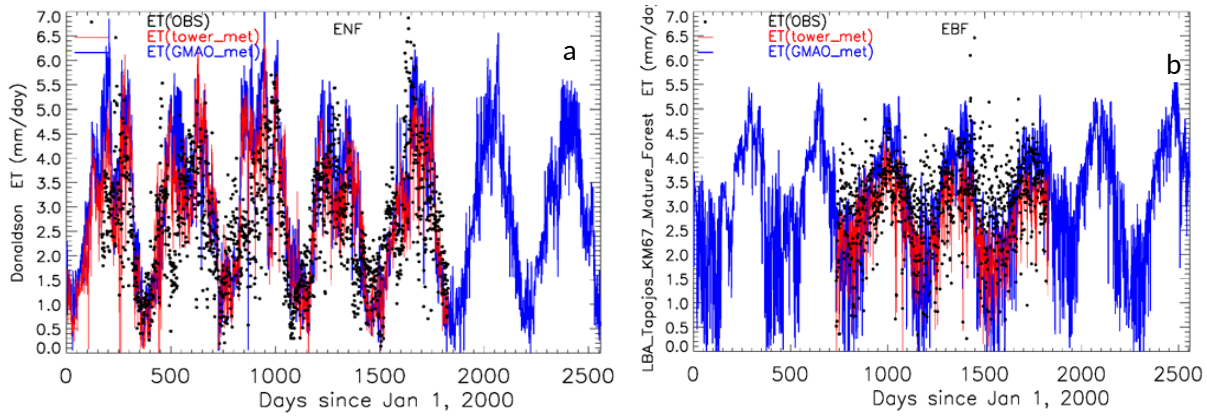


Figure 4.1. The ET measurements (black dots, OBS), the ET estimates driven by flux tower measured meteorological data (red lines) and GMAO meteorological data (blue lines) over 2000-2006 at seven tower sites, Donaldson (a) and LBA Tapajos KM67 Mature Forest (b).

The average daily ET biases between ET observations and ET estimates across the 46 towers are -0.11 kg/m^2/day driven by tower meteorological data and -0.02 kg/m^2/day driven by GMAO meteorological data. The average mean absolute errors (MAE) are 0.33 kg/m^2/day (tower-specific meteorology) and 0.31 kg/m^2/day (GMAO meteorology). The MAE values are 24.6% and 24.1% of the ET measurements, within the 10-30% range of the accuracy of ET observations (Courault et al. 2005; Jiang et al. 2004; Kalma et al. 2008).

5. Practical Details for downloading MOD16 Data

All MODIS land data products are distributed to global users from the USGS Land Processes Distributed Active Archive Center (USGS LP DAAC), found here:

https://lpdaac.usgs.gov/

Specific details about the MODIS land products can be found here:

https://lpdaac.usgs.gov/dataset_discovery/modis

including details about sensor spectral bands, spatial/temporal resolution, platform overpass timing, datafile naming conventions, tiling formats, processing levels and more.

When this document is being written, MODAPS at NASA is testing the operational Collection6 MOD16 code and 500m MOD16 data will be released to the public through the USUS LP DAAC soon.

The long-term consistent improved global Collection5 1-km MOD16 from year of 2000 to the previous year can be downloaded at NTSG at site

http://www.ntsg.umt.edu/project/mod16#data-product

6. MOD16 Data Description and Process

6.1. Description and Process of MOD16 Data Files

There are two major MOD16 data sets, 8-day composite MOD16A2 and annual composite MOD16A3. Both MOD16A2 and MOD16A3 are stored in HDFEOS2 scientific data file format (http://hdfeos.org/software/library.php). HDFEOS2 file format is an extension of HDF4 by adding geo-reference, map projection, and other key meta data information to HDF4 format (https://support.hdfgroup.org/products/hdf4/) to facilitate users to use satellite data products from NASA's Earth Observing System (EOS) projects. Since MOD16 is a level 4 EOS data product, the grid data sets are saved in Sinusoidal (SIN) map projection, an equal-area map projection, with an earth radius of 6371007.181 meters (Note the inversed lat/lon are in WGS84 datum). The MODIS high-level data sets divide the global SIN into many chunks, so-called 10-degree tiles (https://modis-land.gsfc.nasa.gov/MODLAND_grid.html). There are 317 land tiles, and among which, 300 tiles (286 tiles for the Collection5) located within latitude of 60°S and 90°N (90°N for the Collection5) have vegetated land pixels. Therefore, for each 8-day Collection6 MOD16A2 and yearly MOD16A3, there are 300 land tiles globally if there are no missing tiles.

When MODIS updates MOD16 from the Collection5 to Collection6, the spatial resolution has increased from nominal 1-km (926.62543313883 meters) to 500m (463.312716569415 meters), to be consistent with changes in the spatial resolution of a major input to MOD16, the 8-day MOD15A2H.

For users don't know how to handle and process MODIS high-level data products, we suggest users use free or commercial software tools, such as MODIS Reprojection Tool (MRT) (https://lpdaac.usgs.gov/tools/modis_reprojection_tool) , HDF-EOS to GeoTIFF Conversion Tool (HEG) (http://hdfeos.org/software/heg.php), or MODIS toolbox in ArcMap (https://blogs.esri.com/esri/arcgis/2011/03/21/global-evapotranspiration-data-accessible-in-arcmap-thanks-to-modis-toolbox/) to handle MOD16.

6.2. Description of MOD16 Date Sets

6.2.1. MOD16A2

Table 6.1 lists science data sets in the 8-day MOD16A2. ET_500m and potential ET (PET), PET_500m, are the ***summation*** of 8-day total water loss through ET (0.1 kg/m^2/8day), whereas the associated latent heat fluxes and its potential, LE_500m and PLE_500m, are the ***average*** total energy over a unit area for a unit day during the composite 8-day period (10000 J/m^2/day). But be cautious that the last 8-day (MOD16A2.A20??361.*.hdf) of each year is ***not*** 8-day but either 5-day or 6-day depending on normal or leap year.

Table 6.1. *The detailed information on science data sets in MOD16A2*

Data Sets	Meaning	Units	Date Type	Valid Range	Scale Factor
ET_500m	8-day total ET	kg/m^2/8d	int16	-32767 ~ 32760	0.1

LE_500m	8-day average LE	J/m^2/d	int16	-32767 ~ 32760	10000
PET_500m	8-day total PET	kg/m^2/8d	int16	-32767 ~ 32760	0.1
PLE_500m	8-day average PLE	J/m^2/d	int16	-32767 ~ 32760	10000
ET_QC_500m	Quality Control	none	uint8	0 ~ 254	none

All data sets in MOD16A2, except Quality Control (QC) data field, ET_QC_500m, have valid value ranging from -32767 to 32700 and are saved in signed 2-byte short integer (int16). Though data attributes list just one _FillValue: 32767 in the head file of MOD16A2 file, there are, in fact, 7 fill values listed below for non-vegetated pixels, which we didn't calculate ET.

Fill value, out of the earth 32767
Water body 32766
Barren or sparsely vegetated 32765
Permanent snow and ice 32764
Permanent wetland 32763
Urban or Built-up 32762
Unclassified 32761

The QC data layer, ET_QC_500m, directly inherits the QC data field, FparLai_QC, from the corresponding MOD15A2 of the same 8-day. Thus detailed information of bitfields in 8 bitword is the same as that from MOD15A2, as detailed below (a copy of the head file of MOD15A2H).

MOD15A2_FparLai_QC_DOC
FparLai_QC 5 BITFIELDS IN 8 BITWORD
MODLAND_QC START 0 END 0 VALIDS 2
MODLAND_QC 0 = Good Quality (main algorithm with or without saturation)
MODLAND_QC 1 = Other Quality (back-up algorithm or fill value)
SENSOR START 1 END 1 VALIDS 2
SENSOR 0 = Terra
SENSOR 1 = Aqua
DEADDETECTOR START 2 END 2 VALIDS 2
DEADDETECTOR 0 = Detectors apparently fine for up to 50% of channels 1,2
DEADDETECTOR 1 = Dead detectors caused >50% adjacent detector retrieval
CLOUDSTATE START 3 END 4 VALIDS 4 (this inherited from Aggregate_QC bits {0,1} cloud state)
CLOUDSTATE 00 = 0 Significant clouds NOT present (clear)
CLOUDSTATE 01 = 1 Significant clouds WERE present
CLOUDSTATE 10 = 2 Mixed cloud present on pixel
CLOUDSTATE 11 = 3 Cloud state not defined,assumed clear
SCF_QC START 5 END 7 VALIDS 5
SCF_QC 000=0 Main (RT) algorithm used, best result possible (no saturation)
SCF_QC 001=1 Main (RT) algorithm used, saturation occured. Good, very usable.
SCF_QC 010=2 Main algorithm failed due to bad geometry, empirical algorithm used
SCF_QC 011=3 Main algorithm faild due to problems other than geometry, empirical algorithm used
SCF_QC 100=4 Pixel not produced at all, value coudn't be retrieved (possible reasons: bad

L1B data, unusable MOD09GA data)

For non-improved near real-time MOD16A2, we suggest users at least exclude cloud-contaminated cells. For the improved and reprocessed MOD16A2, users may ignore QC data layer because cloud-contaminated LAI/FPAR gaps have been temporally filled before calculating ET (Mu e al., 2007, also see previous section 3.2.1 and following 6.3). QC just denotes if filled LAI/FPAR were used as inputs.

6.2.2. MOD16A3

Table 6.2 lists science data sets in annual MOD16A3. ET_500m and PET_500m are the ***summation*** of total daily ET/PET through the year (0.1 kg/m^2/year) whereas LE and PLE are the corresponding ***average*** total latent energy over a unit area for a unit day (10000 J/m^2/day) through the year. LE_500m and PLE_500m have the same unit, data type (signed 2-byte short int16), valid range and fill values as those listed above for the 8-day MOD16A2; whereas annual ET_500m and PET_500m are saved in unsigned 2-byte short integer (uint16) with valid range from 0 to 65500. Though data attributes list one _FillValue: 65535 in the HDFEOS MOD16A3 file, there are, in fact, 7 fill values as listed below for non-vegetated pixels without ET calculations.

Fill value, out of the earth 65535
Water body 65534
Barren or sparsely vegetated 65533
Permanent snow and ice 65532
Permanent wetland 65531
Urban or Built-up 65530
Unclassified 65529

Table 6.2. *The detailed information on science data sets in MOD16A3*

Data Sets	Meaning	Units	Date Type	Valid Range	Scale Factor
ET_500m	annual sum ET	kg/m^2/yr	uint16	0 ~ 65500	0.1
LE_500m	annual average LE	J/m^2/d	int16	0 ~ 32700	10000
PET_500m	annual sum PET	kg/m^2/yr	uint16	0 ~ 65500	0.1
PLE_500m	annual average PLE	J/m^2/d	int16	0 ~ 32700	10000
ET_QC_500m	Quality Assessment	Percent (%)	uint8	0 ~ 100	none

QC data field in annual MOD16A3, ET_QC_500m, is different from most MODIS QC data sets because it is not bitfields but a more meaningful QC assessment for annual composite values. We used the method proposed by Zhao et al. (2005) to define annual ET QC as

$$ET\_QC\_500m = 100.0 \times NUg/Totalg \quad (26)$$

where NUg is the number of days during growing season with filled MODIS 500m LAI inputs to MOD16 due to missing or unfavorable atmospheric contaminated MODIS LAI (hence FPAR) if improvement reprocess is employed. Totalg is total number of days in the growing season. The

growing season is defined as all days with Tmin above the value where stomata close as in the BPLUT.

The data type of ET_QC_500m is unsigned 1-byte integer (uint8) with valid range from 0 to 100. For vegetated land pixels, if ET_QC_500m has no spatial variations, it will imply the MOD16A3 is the near real-time data product, but not the improved by reprocessing because frequency of cloud contaminations varies with space. Though data attributes list one _FillValue: 255 in the HDFEOS MOD16A3 file, there are, in fact, 7 fill values as listed below for non-vegetated pixels.

Fill value, out of the earth 255
Water body 254
Barren or sparsely vegetated 253
Permanent snow and ice 252
Permanent wetland 251
Urban or Built-up 250
Unclassified 249

6.3. Evaluation of NASA Near Real-time Operational MOD16 with NTSG Gap-filled

The following subsection 1) reveals the differences between NASA near real-time and NTSG gap-filled MOD16 by showing one 8-day global browser of ET in Figure 6.1; 2) demonstrates close global agreements between NASA near real-time Collection6 MOD16A2, MYD16A2 and NTSG Collection5 gap-filled data. Based on our evaluation, we conclude that NASA near real-time Collection6 MOD16A2 and MYD16A3 are ready for users to apply the data to studies of global terrestrial water and energy cycles and environmental changes.

As mentioned in the previous sections, NASA near real-time (NRT) MOD16A2 contains data gaps mainly due to cloudiness or snow cover, which obscures the ground information observed by MODIS and results in unrealistic biophysical variables derived from MODIS, the inputs to MOD16 (Figure 6.1). The reprocessed MOD16 dataset from Numerical Terradynamic Simulation Group (NTSG) at University of Montana (UMT) is a gap-filled dataset which have solved the issue of cloud- or snow-contaminations by temporally gap-filling the biophysical variables using the data from the uncontaminated data prior and post the contaminated periods (Zhao et al., 2005; Mu et al., 2007; Mu et al., 2011).

Figure 6.1 shows an 8-day (2008281, early October) MODIS ET from gap-filled NTSG Collection5 (C5) and NRT Collection (C6) NASA MOD16 and MYD16. It clearly shows large areas with data gaps which are mainly caused by cloudiness or snow cover. Currently, fill value denoting "Barren or sparsely vegetated" are used as value for these data gaps with vegetated pixels in NRT data set which is shown in bronze color in the image. In tropical rainforests, there are large areas of data gaps caused by cloudiness in the NASA NRT MOD16A2 8-day ET datasets, similar phenomena occur to the other three variables: LE, PET and PLE (not shown). The main reason for more data gaps in MYD16 than in MOD16 is the difference in the local overpass time between TERRA and Aqua satellite. TERRA's overpass time is in the morning whereas Aqua in the afternoon, and convection is much stronger in the afternoon when the land

surface is warmer than in the morning when it is cooler. Stronger convection potentially can induce more cloudiness.

We evaluated the NASA C6 NRT MOD16 by comparing it with gap-filled NTSG C5 just for commonly valid pixels from both data sets. Because C6 and C5 have different versions of MODIS inputs, such as MODIS land cover, LAI/FPAR, and surface albedo, and the two also use different versions of daily meteorological data sets (operational GMAO for NRT and long-term consistent MERRA/GMAO for NTSG), it is expected that the two data sets would have differences. In addition, the C6 has a higher spatial resolution (500m) then the C5 (1-km). We randomly chose year 2008 because it is after year 2002, when a full yearly MODIS datasets are available from either TERRA or Aqua. To perform pixel-by-pixel comparisons, we first smoothed the NASA C6 NRT 500m data into 1-km.

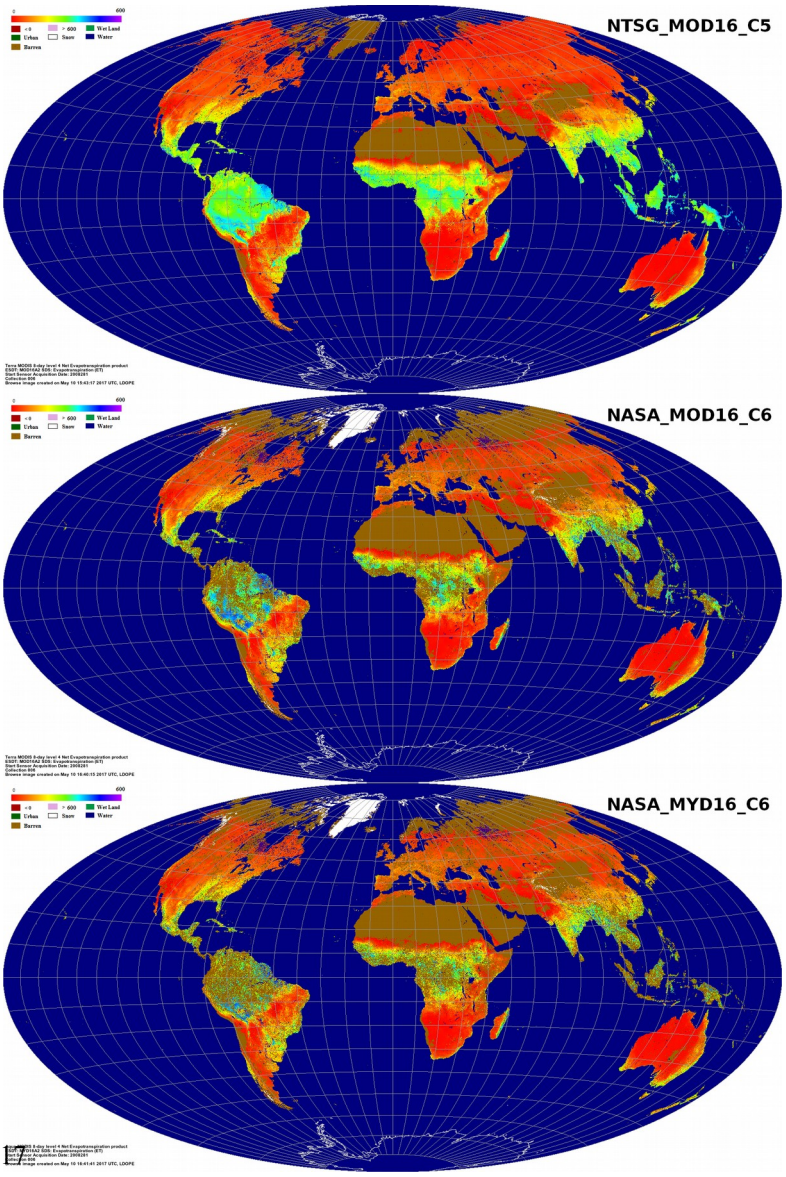


Figure 6.1. *Comparison of MOD16A2 for an 8-day of 2008281 (Oct 7th through Oct 14th) between NTSG Collection5 gap-filled and NASA near real-time Collection6 with data gaps due to cloudiness or snow cover.*

We used NTSG gap-filled MOD16 as baseline for the evaluation. We compared all 46 8-day ET, LE, PET and PLE for both MOD16A2 and MYD16A2 and three statistic metrics (mean, STD, and correlation) are used. Mean and TSD can reveal if the two data sets have similar magnitude and range of variations, and correlation can reveal if the two data sets have similar directions of variations. Though there is no gap-filled MYD16A2 from NTSG, we still use NTSG MOD16A2 as baseline to assess NASA NRT MYD16A2. We also separated MOD from MYD to have two sets of evaluations in order to maximize the number of valid pixels involved for the assessments. Otherwise, as shown in Figure 6.1, there would be much fewer common valid pixels if we put NASA MOD16 and MYD16 together.

Figure 6.2 shows the seasonality of mean, STD, and correlation for ET, LE, PET and PLE between baseline NTSG MOD16A2 and NASA NRT MOD16A2; and between baseline NTSG MOD16A2 and NASA NRT MYD16A2, respectively. Correspondingly, Table 6.3 shows the averages of difference ($\acute{Diff}$) in mean, relative differences ($\acute{R}_{Diff}$) in %, and average correlation ($\acute{Cor}$) across all 46 8-days in 2008 between the NASA NRT MOD16A2 and MYD16A2 and NTSG MOD16A2, respectively. $\acute{Diff}$ and $\acute{R}_{Diff}$ are defined below,

$$\acute{Diff} = \frac{\sum_{i=1}^{46} (NASA\ NRT - NTSG)}{46} \tag{27}$$

$$\acute{R}_{Diff} = \frac{100 * \acute{Diff}}{\acute{NTSG}} \tag{28}$$

Table 6.3. A) Averages of difference ($\acute{Diff}$), relative differences ($\acute{R}_{Diff}$) in %, and average correlation ($\acute{Cor}$) across all 46 8-days in 2008 between the NASA NRT and NTSG MOD16A2 by four latitudinal zones.

MOD16A2	**ET**		**LE**		**PET**		**PLE**	
Latitudinal Zones	$\acute{Diff}$, $\acute{R}_{Diff}$	$\acute{Cor}$	$\acute{Diff}$, $\acute{R}_{Diff}$	$\acute{Cor}$	$\acute{Diff}$, $\acute{R}_{Diff}$	$\acute{Cor}$	$\acute{Diff}$, $\acute{R}_{Diff}$	$\acute{Cor}$
47.5°N~80°N	0.65, 7.06%	0.86	0.20, 6.88%	0.86	-0.57, -2.73%	0.93	-0.32, -4.97%	0.92
22.5°N~47.5°N	0.49, 5.06%	0.91	0.14, 4.76%	0.92	0.90, 2.33%	0.91	0.24, 2.06%	0.91
-22.5°S~22.5°N	-0.12,	0.90	-0.06,	0.90	2.49,		0.69,	

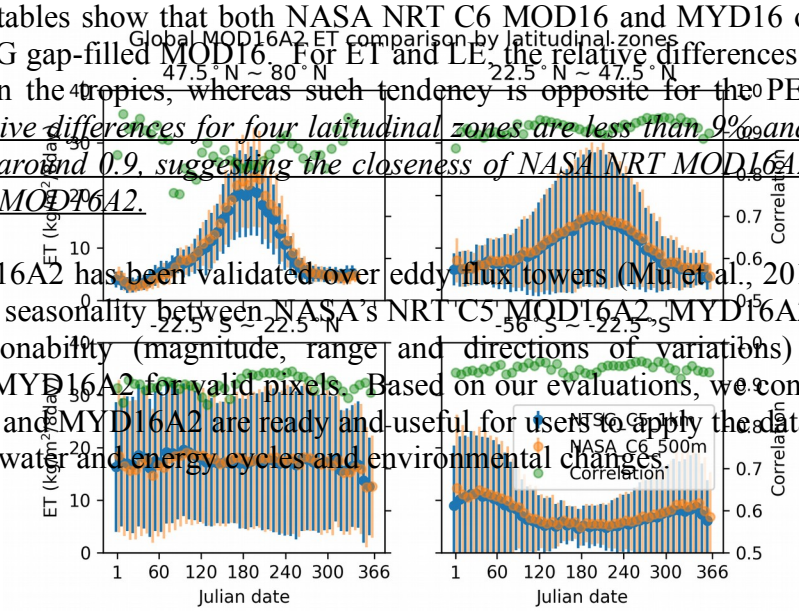
	-0.67%		-1.05%		5.02%	0.77	4.55%	0.77
-56°S~-22.5°S	0.56, 7.90%	0.94	0.16, 7.14%	0.934	0.49, 1.00%	0.93	0.05, 0.33%	0.94

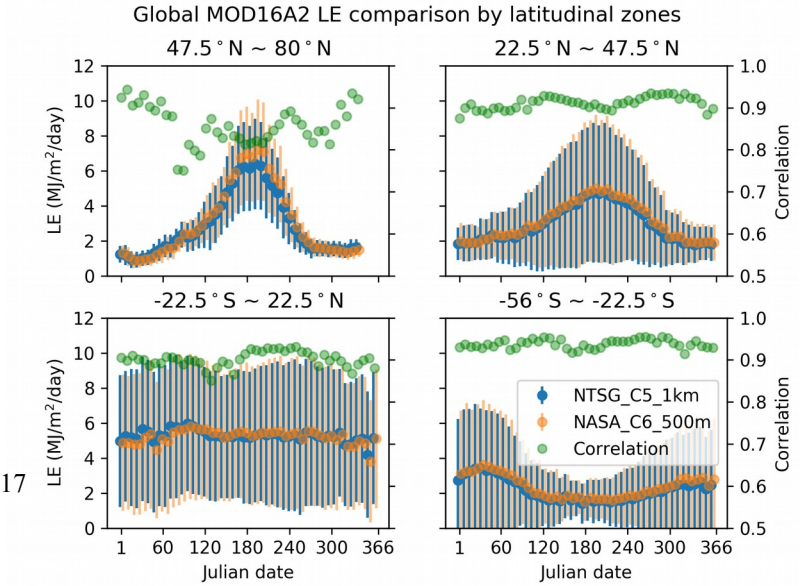
Table 6.3. B) Similar to A) but for NASA NRT MYD16A2 against NTSG MOD16

MYD16A2	**ET**		**LE**		**PET**		**PLE**	
Latitudinal Zones	$Diff$, R_{Diff}	Cor	$Diff$, R_{Diff}	Cor	$Diff$, R_{Diff}	Cor	$Diff$, R_{Diff}	Cor
47.5°N~80°N	0.66, 7.16%	0.84	0.19, 6.74%	0.84	-0.52, -2.45%	0.93	-0.32, -4.83%	0.92
22.5°N~47.5°N	0.48, 5.08%	0.90	0.12, 4.21%	0.90	1.25, 3.21%	0.91	0.28, 2.33%	0.91
-22.5°S~22.5°N	-0.21, -1.24%	0.88	-0.11, -2.11%	0.88	2.88, 5.83%	0.77	0.69, 4.57%	0.78
-56°S~-22.5°S	0.59, 8.43%	0.93	0.14, 6.62%	0.93	1.00, 2.04%	0.93	0.03, 0.21%	0.94

The figures and tables show that both NASA NRT C6 MOD16 and MYD16 overall tend to be higher than NTSG gap-filled MOD16. For ET and LE, the relative differences are greater in the extra tropics than the tropics, whereas such tendency is opposite for the PET and PLE. *The averages of relative differences for four latitudinal zones are less than 9% and averages of the correlations are around 0.9, suggesting the closeness of NASA NRT MOD16A2 to the baseline, NTSG gap-filled MOD16A2.*

NTSG C5 MOD16A2 has been validated over eddy flux towers (Mu et al., 2011), and the close agreement in the seasonality between NASA's NRT C5 MOD16A2, MYD16A2 and NTSG data reveals the reasonability (magnitude, range and directions of variations) of NASA NRT MOD16A2 and MYD16A2 for valid pixels. Based on our evaluations, we conclude that NASA NRT MOD16A2 and MYD16A2 are ready and useful for users to apply the data to the studies of global terrestrial water and energy cycles and environmental changes.





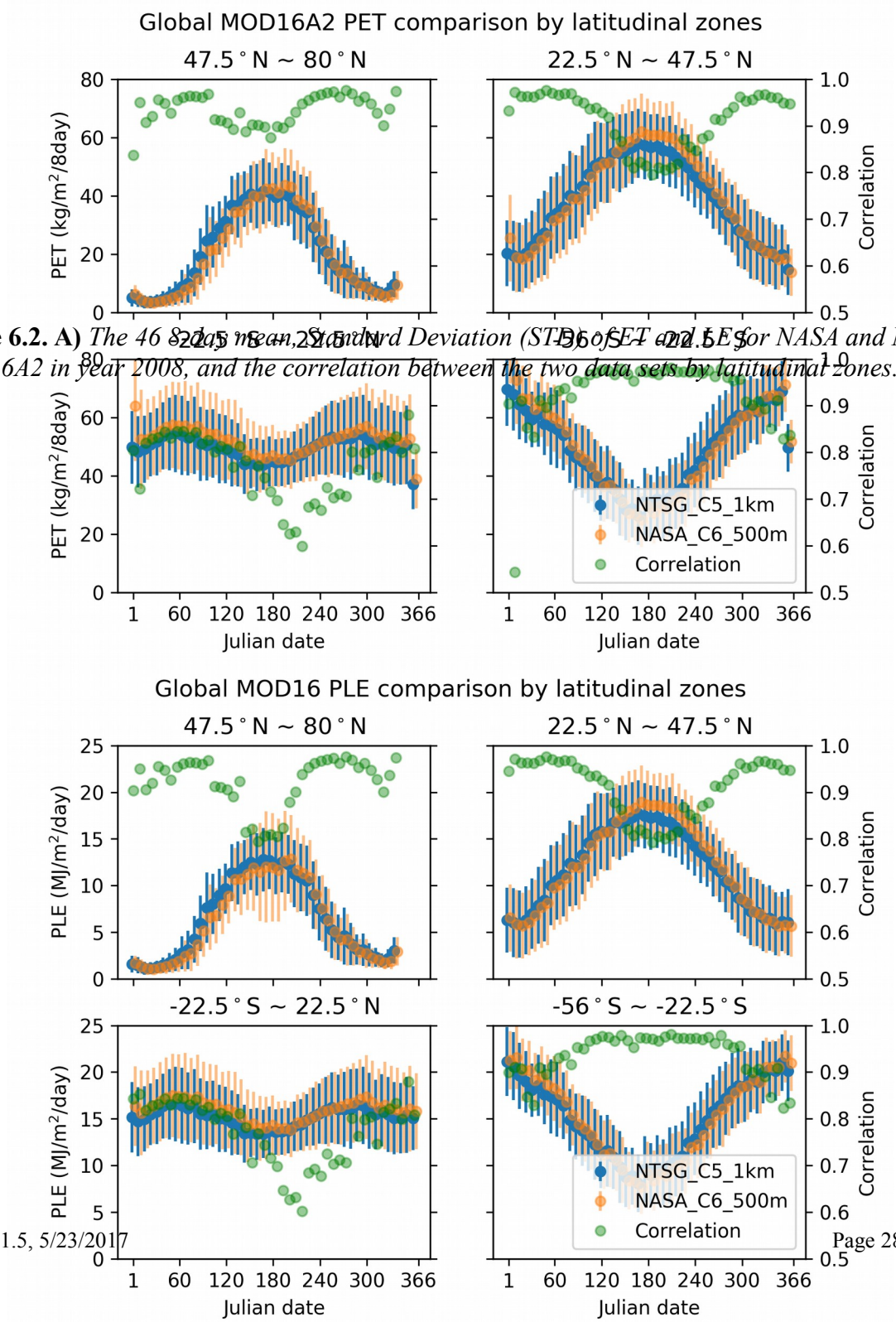


Figure 6.2. A) *The 46 8-day mean, Standard Deviation (STD) of ET and LE for NASA and NTSG MOD16A2 in year 2008, and the correlation between the two data sets by latitudinal zones.*

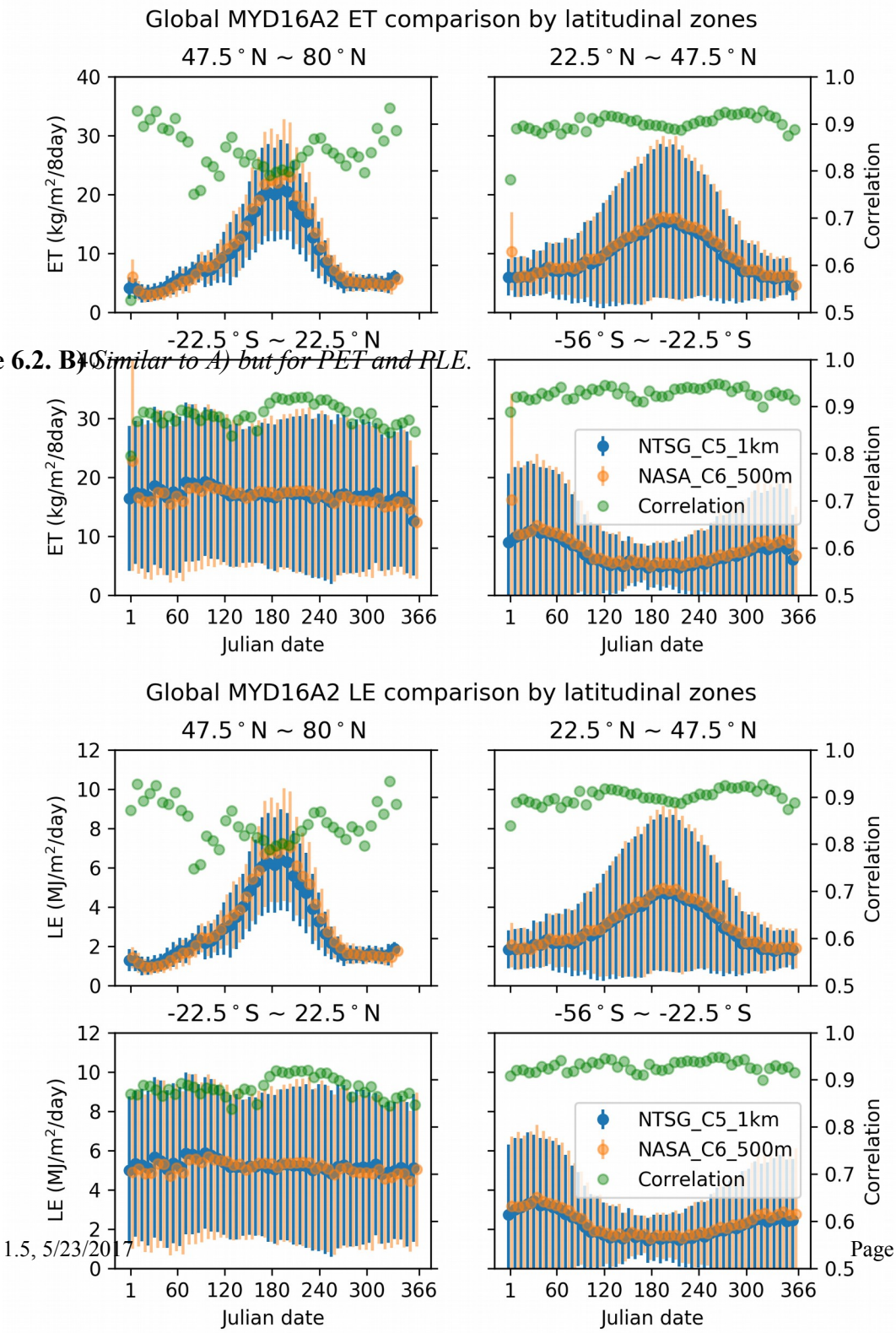


Figure 6.2. B) *Similar to A) but for PET and PLE.*

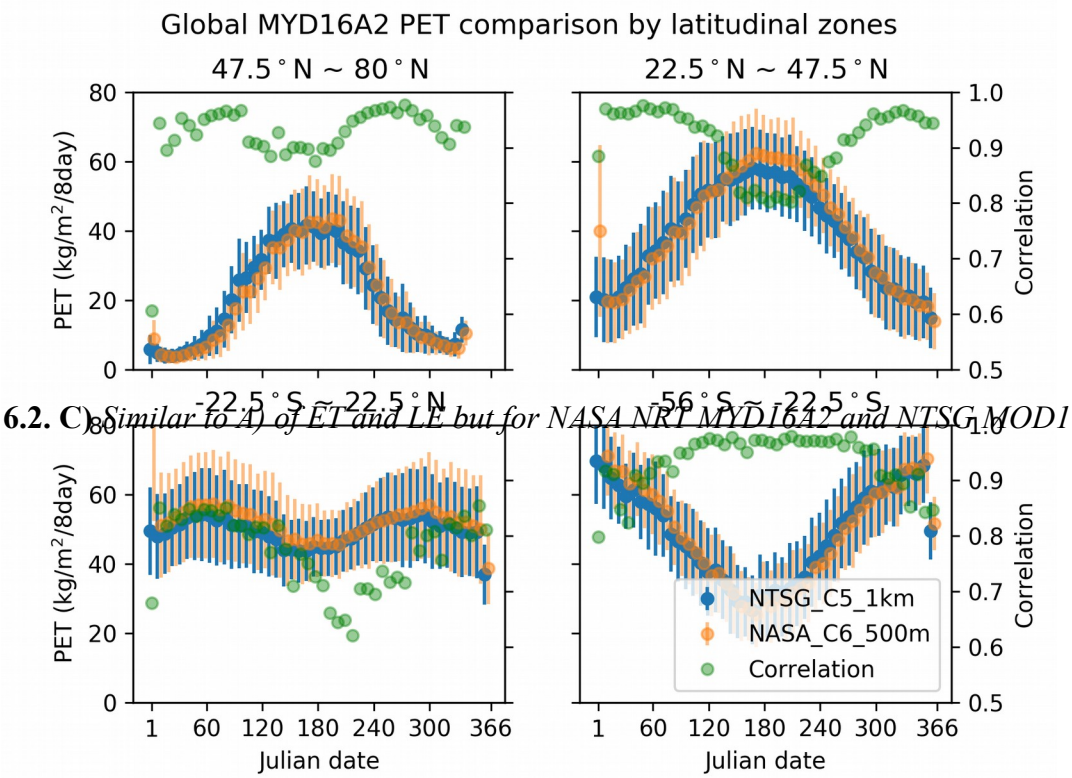


Figure 6.2. C) *Similar to A) of ET and LE but for NASA NRT MYD16A2 and NTSG MOD16A2.*

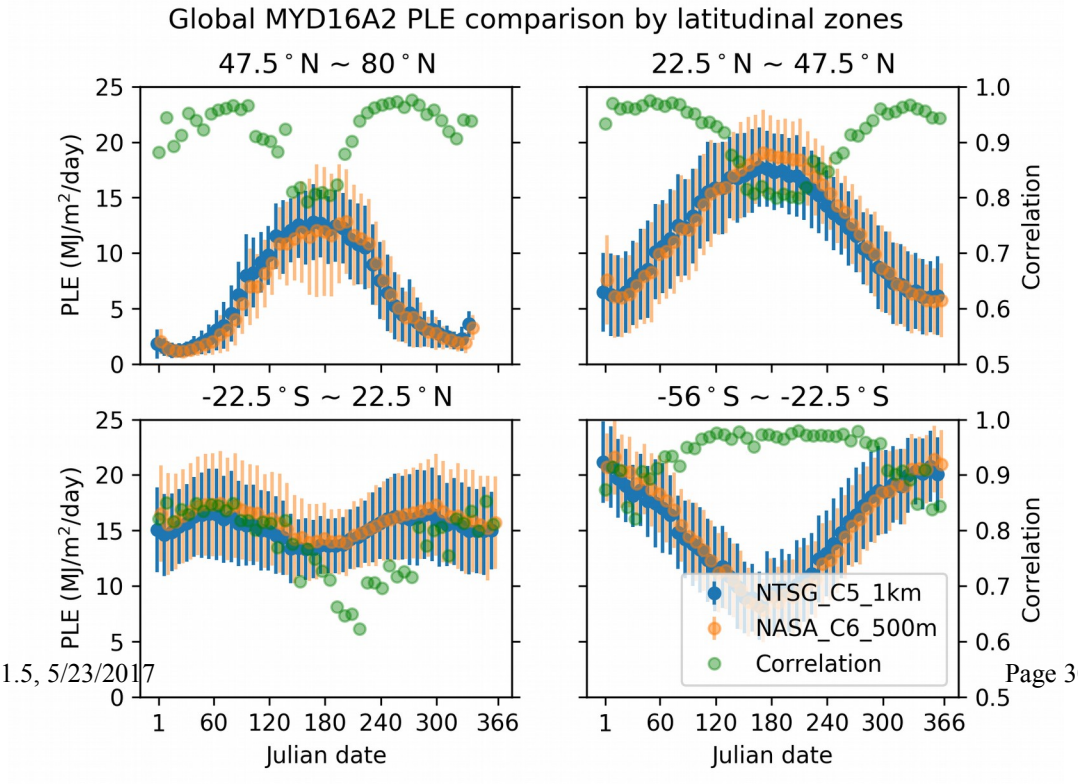


Figure 6.2. D) *Similar to A) but for PET and PLE of NASA NRT MYD16A2 and NTSG MOD16A2.*

LIST OF NTSG AUTHORED/CO-AUTHORED PAPERS USING MOD16 ET: 2007 – 2015

[all available at http://www.ntsg.umt.edu/biblio]

Chen, Y., J. Xia, S. Liang, J. Feng, J. B. Fisher, X. Li, X. Li et al. (2014) Comparison of satellite-based evapotranspiration models over terrestrial ecosystems in China. *Remote Sensing of Environment* 140 (2014): 279-293.

Cleugh, H. A., Leuning, R., Mu, Q., and Running, S. W. (2007). Regional evaporation estimates from flux tower and MODIS satellite data. *Remote Sensing of Environment*, *106*(3), 285-304.

Jovanovic, N., Mu, Q., Bugan, R. D., and Zhao, M. (2015). Dynamics of MODIS evapotranspiration in South Africa. *Water SA*, *41*(1), 79-90.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., ... and Dolman, A. J. (2010). Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, *467*(7318), 951-954.

Mu, Q., Heinsch, F. A., Zhao, M., and Running, S. W. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sensing of Environment*, *111*(4), 519-536.

Mu, Q., Jones, L. A., Kimball, J. S., McDonald, K. C., and Running, S. W. (2009). Satellite assessment of land surface evapotranspiration for the pan-Arctic domain. *Water Resources Research*, *45*(9).

Mu, Q., Zhao, M., and Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, *115*(8), 1781-1800.

Mu, Q., Zhao, M., and Running, S. W. (2011). Evolution of hydrological and carbon cycles under a changing climate. *Hydrological Processes*, *25*(26), 4093-4102.

Mu, Q., Zhao, M., & Running, S. W. (2012). Remote Sensing and Modeling of Global Evapotranspiration. In *Multiscale Hydrologic Remote Sensing: Perspectives and Applications* (pp. 443-480), edited by Chang, N. B., Hong, Y., CRC Press.

Mu, Q., Zhao M., Kimball J. S., McDowell N., and Running S. W., (2013) A Remotely Sensed Global Terrestrial Drought Severity Index, *Bulletin of the American Meteorological Society*, 94: 83–98.

Renzullo, L. J., Barrett, D. J., Marks, A. S., Hill, M. J., Guerschman, J. P., Mu, Q., and Running, S. W. (2008). Multi-sensor model-data fusion for estimation of hydrologic and energy flux parameters. *Remote Sensing of Environment*, *112*(4), 1306-1319.

Ruhoff, A. L., Paz, A. R., Aragao, L. E. O. C., Mu, Q., Malhi, Y., Collischonn, W., ... and Running, S. W. (2013). Assessment of the MODIS global evapotranspiration algorithm using eddy covariance measurements and hydrological modelling in the Rio Grande basin. *Hydrological Sciences Journal*, *58*(8), 1658-1676.

Smettem, K. R., Waring, R. H., Callow, J. N., Wilson, M., and Mu, Q. (2013). Satellite‐derived estimates of forest leaf area index in southwest Western Australia are not tightly coupled to interannual variations in rainfall: implications for groundwater decline in a drying climate. *Global Change Biology*, *19*(8), 2401-2412.

Wang, S., Pan, M., Mu, Q., Shi, X., Mao, J., Brümmer, C., ... and Black, T. A. (2015). Comparing Evapotranspiration from Eddy Covariance Measurements, Water Budgets, Remote Sensing, and Land Surface Models over Canada. *Journal of Hydrometeorology*, *16*(4), 1540-1560.

Yao, Y., Liang, S., Li, X., Hong, Y., Fisher, J. B., Zhang, N., ... and Jiang, B. (2014). Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations. *Journal of Geophysical Research: Atmospheres*, *119*(8), 4521-4545.

Zhang, K., Kimball, J. S., Mu, Q., Jones, L. A., Goetz, S. J., and Running, S. W. (2009). Satellite based analysis of northern ET trends and associated changes in the regional water balance from 1983 to 2005. *Journal of Hydrology*, *379*(1), 92-110.

REFERENCES

Belward, A. S., Estes, J. E., and Kline, K. D. (1999). The IGBP-DIS global 1-km land-cover data set DISCover: A project overview. *Photogrammetric Engineering and Remote Sensing*, *65*(9), 1013-1020.

Bouchet, R. J. (1963). Evapotranspiration reelle, evapotranspiration potentielle, et production agricole. *Ann. agron*, *14*(5), 743-824.

Choudhury, B. (2000). A Biophysical Process-Based Estimate of Global Land Surface Evaporation Using Satellite and Ancillary Data. In *Observing Land from Space: Science, Customers and Technology* (pp. 119-126). Springer Netherlands.

Cleugh, H. A., Leuning, R., Mu, Q., and Running, S. W. (2007). Regional evaporation estimates from flux tower and MODIS satellite data. *Remote Sensing of Environment*, *106*(3), 285-304.

Courault, D., Seguin, B., and Olioso, A. (2005). Review on estimation of evapotranspiration from remote sensing data: From empirical to numerical modeling approaches. *Irrigation and Drainage systems*, *19*(3-4), 223-249.

Dang, Q. L., Margolis, H. A., Coyea, M. R., Sy, M., and Collatz, G. J. (1997). Regulation of branch-level gas exchange of boreal trees: roles of shoot water potential and vapor pressure difference. *Tree Physiology*, *17*(8-9), 521-535.

Dawson, T. E., Burgess, S. S., Tu, K. P., Oliveira, R. S., Santiago, L. S., Fisher, J. B., ... and Ambrose, A. R. (2007). Nighttime transpiration in woody plants from contrasting ecosystems. *Tree Physiology*, *27*(4), 561-575.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*(1), 168-182.

Fisher, J. B., Tu, K. P., and Baldocchi, D. D. (2008). Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. *Remote Sensing of Environment*, *112*(3), 901-919.

Jarvis, P. G. (1976). The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*,*273*(927), 593-610.

Jiang, L., Islam, S., and Carlson, T. N. (2004). Uncertainties in latent heat flux measurement and estimation: implications for using a simplified approach with remote sensing data. *Canadian Journal of Remote Sensing*, *30*(5), 769-787.

Jones, H. G. (1992). *Plants and microclimate: a quantitative approach to environmental plant physiology*. Cambridge university press.

Kalma, J. D., McVicar, T. R., and McCabe, M. F. (2008). Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. *Surveys in Geophysics*, *29*(4-5), 421-469.

Kawamitsu, Y., Yoda, S., and Agata, W. (1993). Humidity pretreatment affects the responses of stomata and CO2 assimilation to vapor pressure difference in C3 and C4 plants. *Plant and Cell Physiology*, *34*(1), 113-119.

Kelliher, F. M., Leuning, R., Raupach, M. R., and Schulze, E. D. (1995). Maximum conductances for evaporation from global vegetation types.*Agricultural and Forest Meteorology*, *73*(1), 1-16.

Leuning, R. (1995). A critical appraisal of a combined stomatal-photosynthesis model for C3 plants. *Plant, Cell & Environment*, *18*(4), 339-355.

Los, S. O., Pollack, N. H., Parris, M. T., Collatz, G. J., Tucker, C. J., Sellers, P. J., ... and Dazlich, D. A. (2000). A global 9-yr biophysical land surface dataset from NOAA AVHRR data. *Journal of Hydrometeorology*,*1*(2), 183-199.

Marsden, B. J., Lieffers, V. J., and Zwiazek, J. J. (1996). The effect of humidity on photosynthesis and water relations of white spruce seedlings during the early establishment phase. *Canadian Journal of Forest Research*,*26*(6), 1015-1021.

Misson, L., Panek, J. A., and Goldstein, A. H. (2004). A comparison of three approaches to modeling leaf gas exchange in annually drought-stressed ponderosa pine forests. *Tree Physiology*, *24*(5), 529-541.

Monteith, J. L. (1965). Evaporation and environment. In *Symposia of the Society for Experimental Biology* (Vol. 19, p. 205).

Mu, Q., Heinsch, F. A., Zhao, M., and Running, S. W. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sensing of Environment*, *111*(4), 519-536.

Mu, Q., Zhao, M., and Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, *115*(8), 1781-1800.

Oren, R., Sperry, J. S., Katul, G. G., Pataki, D. E., Ewers, B. E., Phillips, N., and Schäfer, K. V. R. (1999). Survey and synthesis of intra-and interspecific variation in stomatal sensitivity to vapour pressure deficit.*Plant, Cell & Environment*, *22*(12), 1515-1526.

Running, S. W., Loveland, T. R., and Pierce, L. L. (1994). A vegetation classification logic based on remote sensing for use in global scale biogeochemical models. *Ambio, 23*, 77-81.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., and Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *Bioscience*, *54*(6), 547-560.

Running, S. W., and Kimball, J. S. (2005). Satellite-Based Analysis of Ecological Controls for Land-Surface Evaporation Resistance. *Encyclopedia of hydrological sciences*.

Sandford, A. P., and Jarvis, P. G. (1986). Stomatal responses to humidity in selected conifers. *Tree Physiology*, *2*(1-2-3), 89-103.

Schubert, S. D., Rood, R. B., and Pfaendtner, J. (1993). An assimilated dataset for earth science applications. *Bulletin of the American meteorological Society*, *74*(12), 2331-2342.

Schulze, E. D., Kelliher, F. M., Korner, C., Lloyd, J., and Leuning, R. (1994). Relationships among maximum stomatal conductance, ecosystem surface conductance, carbon assimilation rate, and plant nitrogen nutrition: a global ecology scaling exercise. *Annual Review of Ecology and Systematics*, 629-660.

Shuttleworth, W. J., and Wallace, J. S. (1985). Evaporation from sparse crops-an energy combination theory. *Quarterly Journal of the Royal Meteorological Society*, *111*(469), 839-855.

Thornton, P.E. (1998). Regional ecosystem simulation: combining surface- and satellite-based observations to study linkages between terrestrial energy and mass budgets. PhD. Dissertation, School of Forestry, *The University of Montana*, Missoula, MT., 280 pp.

Tuzet, A., Perrier, A., and Leuning, R. (2003). Stomatal control of photosynthesis and transpiration: Results from a soil–plant–atmosphere continuum model. *Plant, Cell and Environment*, *26*, 1097-1116.

van de Griend, A. A., and Owe, M. (1994). Bare soil surface resistance to evaporation by vapor diffusion under semiarid conditions. *Water Resources Research*, *30*, 181-188.

Wan, Z., Zhang, Y., Zhang, Q., and Li, Z. L. (2002). Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. *Remote sensing of Environment*, *83*(1), 163-180.

White, M. A., Thornton, P. E., Running, S. W., and Nemani, R. R. (2000). Parameterization and sensitivity analysis of the BIOME-BGC terrestrial ecosystem model: net primary production controls. *Earth interactions*, *4*(3), 1-85.

Xu, L., and Baldocchi, D. D. (2003). Seasonal trends in photosynthetic parameters and stomatal conductance of blue oak (Quercus douglasii) under prolonged summer drought and high temperature. *Tree Physiology*, *23*(13), 865-877.Zhao, M., Heinsch, F. A., Nemani, R. R., and Running, S. W. (2005). Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment*, *95*(2): 164–176.

Zhao, M., Running, S. W., and Nemani, R. R. (2006). Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses. *Journal of Geophysical Research*, *111, G01002*, doi:10.1029/2004JG000004